

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G06F 17/30	A1	(11) International Publication Number: WO 99/39282 (43) International Publication Date: 5 August 1999 (05.08.99)
<p>(21) International Application Number: PCT/JP99/00259</p> <p>(22) International Filing Date: 22 January 1999 (22.01.99)</p> <p>(30) Priority Data: 9801784.1 29 January 1998 (29.01.98) GB</p> <p>(71) Applicant (for all designated States except US): SHARP KABUSHIKI KAISHA [JP/JP]; 22-22, Nagaike-cho, Abeno-ku, Osaka-shi, Osaka 545-8522 (JP).</p> <p>(72) Inventor; and (75) Inventor/Applicant (for US only): SANFILIPPO, Antonio, Pietro [IT/LU]; 18a, rue des Aubepines, L-1145 Luxembourg (LU).</p> <p>(74) Agent: YAMAMOTO, Shusaku; Crystal Tower, 15th floor, 2-27, Shiroimi 1-chome, Chuo-ku, Osaka-shi, Osaka 540-6015 (JP).</p>		<p>(81) Designated States: CA, CN, IN, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</p> <p>Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p>
<p>(54) Title: SCORING OF TEXT UNITS</p> <p>(57) Abstract</p> <p>A method of operating on a text comprising a plurality of text units, each text unit comprising one or more strings, so as to rank said text units, comprises: forming a structure for each of at least some of said strings, in which structure a string is associated with each text unit in which the string occurs; for each text unit summing the number of occurrences of each other text unit in said structures so as to form an individual score for each pair of text units; and processing the individual scores for each text unit in order to form a final score for each text unit on the basis of which the text units can be ranked.</p>		